DOCUMENT RESUME

ED 459 619 FL 027 106

AUTHOR Epstein, Melissa A.; Ladefoged, Peter
TITLE Phonation Types in Amerindian Languages.
INSTITUTION Alaska Univ., Fairbanks. Alaska Native Language Center.
PUB DATE 2001-05-00
NOTE 11p.; Paper presented at the Annual Meeting of the Athabaskan Languages Conference (Los Angeles, CA, May 18-20, 2001). In: Tuttle, S. and G. Holten, Eds. "Proceedings of the 2001 Athabaskan Languages Conference (Los Angeles, May 18-20, 2001)" Working Paper No. 1. ANLC (2001) pp. 154-162.
AVAILABLE FROM Alaska Native Language Center, University of Alaska Fairbanks, Box 757680, Fairbanks, AK 99775-7680 (papers not sold individually; proceedings, $10 plus $3.95 shipping). Tel: 907-474-7874; Web site: http://www.uaf.edu/ancl/index.html.
PUB TYPE Reports - Evaluative (142) -- Speeches/Meeting Papers (150)
EDRS PRICE EDRS Price MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Acoustic Phonetics; *Athapascan Languages; Distinctive Features (Language); Foreign Countries; Grammar; Intonation; *Phonology; Pronunciation; Salish; Second Language Instruction; Second Language Learning; Speech Communication; *Structural Analysis (Linguistics); Syntax; Uncommonly Taught Languages
IDENTIFIERS Mazatec; Mexico; Salish (Tribe); Zapotec

ABSTRACT

This paper lays out techniques for investing different types of phonation from acoustic data. Nearly all the world's languages distinguish between voiced and voiceless phonation. Many languages use other phonation types, such as the breathy voice--also described as speaking while sighing. Another kind of voice quality, one frequently found in Athabaskan languages, is creaky voice in which the vocal folds are pressed together more tightly than usual. So-called glottalized sounds often have this voice quality. Creaky and breathy voice differ from regular voice--called modal voice--in opposite ways. In breathy voice, the vocal chords are further apart and let more air pass through than in modal voice, whereas in creaky voice (also called pressed voice) the vocal chords are pressed tightly together, largely blocking the airflow. Acoustic data analyzed by computer is the means of gathering data for this research project. Specifically, this paper looks at the breathy voice of Jalapa Mazatec, an Otomanguen language spoken in Mexico. This is contrasted with the squeaky voice of Montana Salish, or San Lucas Quiavini Zapotec. (KFT)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 459 619

# Phonation types in Amerindian languages

Melissa A. Epstein and Peter Ladefoged
Linguistics Department, UCLA

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY

M. Krause

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2A

FL027106



BEST COPY AVAILABLE

## Phonation types in Amerindian languages

Melissa A. Epstein and Peter Ladefoged
Linguistics Department, UCLA

This paper lays out techniques for investigating different types of phonation from acoustic data. Nearly all of the world's languages distinguish between voiced and voiceless phonation. Many languages also use other phonation types, such as breathy voice, in which the vocal folds are vibrating, but loosely, so that they allow a considerable amount of breath to pass between them. This kind of voice has also been described as speaking while sighing. Another kind of voice quality, one frequently found in Athabaskan languages, is creaky voice in which the vocal folds are pressed more tightly together than usual. So-called glottalized sounds often have this voice quality. Creaky voice and breathy voice differ from regular voice (which we will call modal voice), in opposite ways. In breathy voice the vocal folds are further apart and let more air through than in modal voice, whereas in creaky voice (which is also called pressed voice) the vocal folds are pressed tightly together, largely blocking the airflow.

The most direct way to investigate different types of phonation would be to look at the vocal folds — an impossible procedure in most linguistic investigations, as it requires medical supervision. Fortunately, however, it is possible to deduce a considerable amount about the state of the glottis by reference to acoustic data. Using a computer, it is possible to analyze recordings and find out how the vocal folds must have been vibrating.

We will begin by looking at breathy voice in Jalapa Mazatec, an Otomanguean language spoken in Mexico. When a recording of the Mazatec word **ndæ̤** 'horse' has been transferred onto a computer we can see various characteristics of the breathy voiced vowel. Figure 1 shows a 150 ms section of the waveform shortly after the release of the stop. The first 50 ms of this section is in the upper part of the figure, the second in the middle, and the third in the lower part. The three parts are slightly overlapped so as to make it clear that this is a continuous waveform. In the first 50 ms (the upper part of the figure) the waveform is very irregular. This is because the vocal folds are apart and the airstream is being set into random variations, much like that of the wind rushing around a corner. The next section begins in this irregular way, but becomes more regular towards the end. At this time the vocal folds are approaching each other and beginning to vibrate without making a complete closure. The final section has six cycles of a repetitive waveform. Each of these cycles is produced by a glottal pulse — a sharp variation in air pressure produced when the vocal folds come together to form a complete closure. Within this display of 150 ms of the waveform we can see that in a phonologically breathy voiced vowel the degree of breathy voicing varies from a great deal of breath and very little voice at the beginning, through a stage when the vocal folds are beginning to vibrate producing a kind of murmured voice, to a fully voiced sound in which the vocal folds are vibrating so as to produce regular glottal pulses.

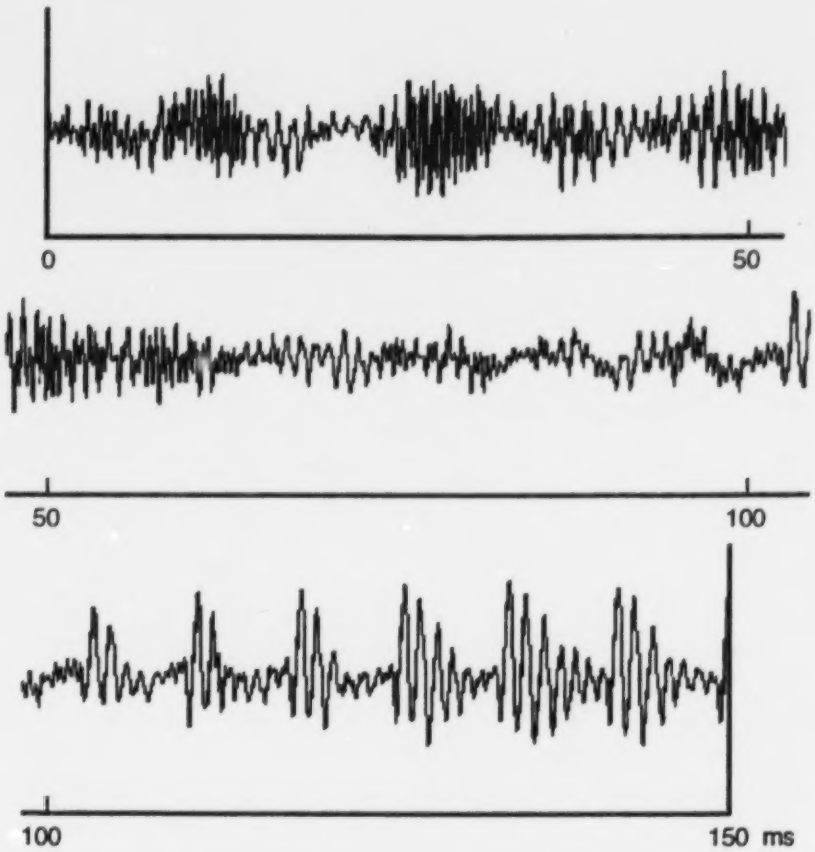



Figure 1. A section of the waveform of the Jalapa Mazatec word **ndæ̰** 'horse'.

Computer displays of the waveform also enable us to see some characteristics of creaky voice. Jalapa Mazatec has creaky vowels as well as modal and breathy vowels, so there is a three-way contrast. Figure 2 shows part of a creaky voiced vowel as produced by five different speakers. The main characteristic of creaky voice that can be seen in the waveform is the irregularity of the interval between consecutive glottal pulses. We can see that the vocal fold pulses are not at exactly regular intervals of time for each of the five speakers in figure 2. Some speakers (e.g. speaker 1) have much more regular vocal fold vibrations than others (e.g. speaker 4), but in all cases the creaky voiced vowel has some degree of irregularity in the intervals between glottal pulses. The technical term for this irregularity is "jitter".

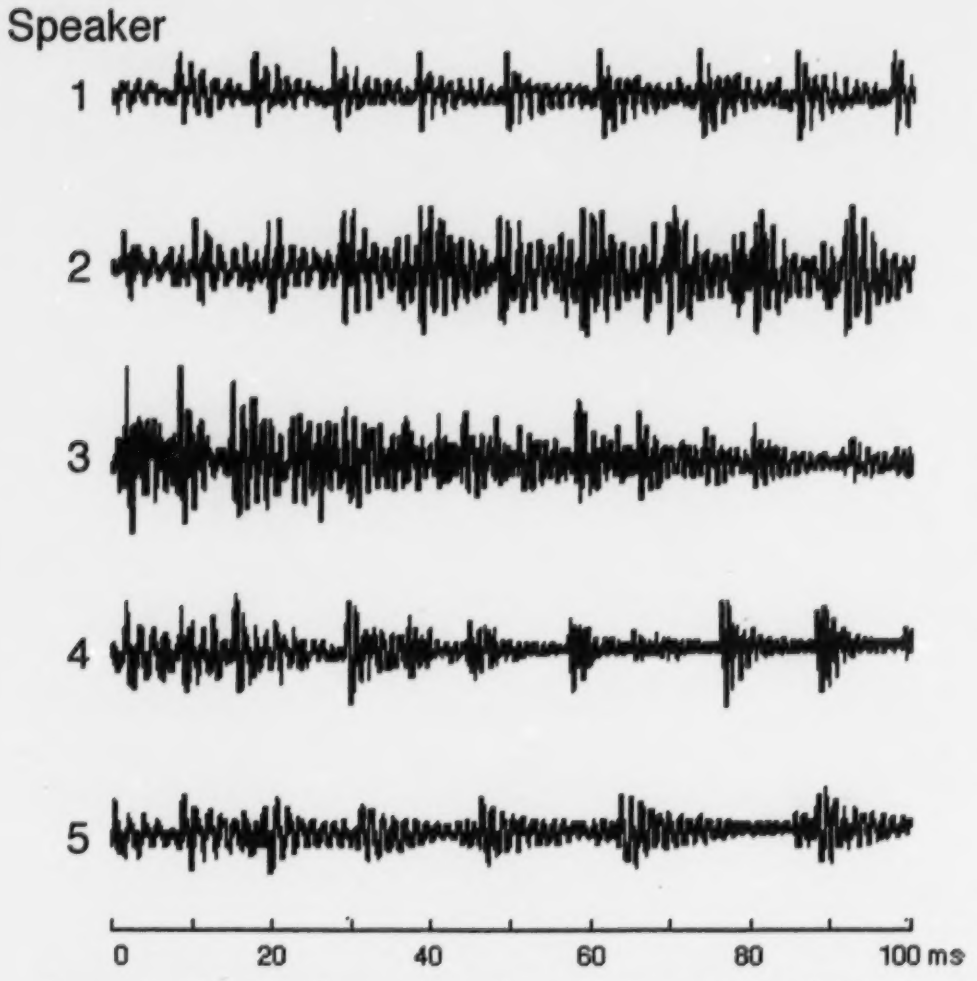



Figure 2. The waveforms of creaky voiced vowels produced by five speakers of Jalapa Mazatec.

We can quantify the degree of jitter in a vowel by using a computer program that enables us to measure the interval between adjacent glottal pulses. Using such a program we can measure the time in milliseconds between each of the glottal pulses for each of the speakers in figure 2. We can then find the average interval between pulses for each speaker. The next step is to see how much each speaker varies from his average. Speakers 1 and 2 vary very little from their average intervals in comparison with speakers 3, 4 and 5. To be more precise, we can take the set of glottal pulse intervals for each speaker and calculate (or let a computer calculate) not only the mean interval but also the standard deviation from the mean.

Marking the individual glottal pulses is sometimes not straightforward. It is possible to interpret the waveform of speaker 3 in different ways, particularly in the

interval between 30 and 50 ms. It is always good scientific practice to try to disprove one's own hypothesis. If we try to make the measurements in the way that is most likely to disprove the hypothesis that creaky voice has irregular pulses, then the pulses for speaker 3 can be marked as shown by the solid arrows in figure 3. Open arrows mark other possible pulse onsets, which would have led to this vowel having more jitter.

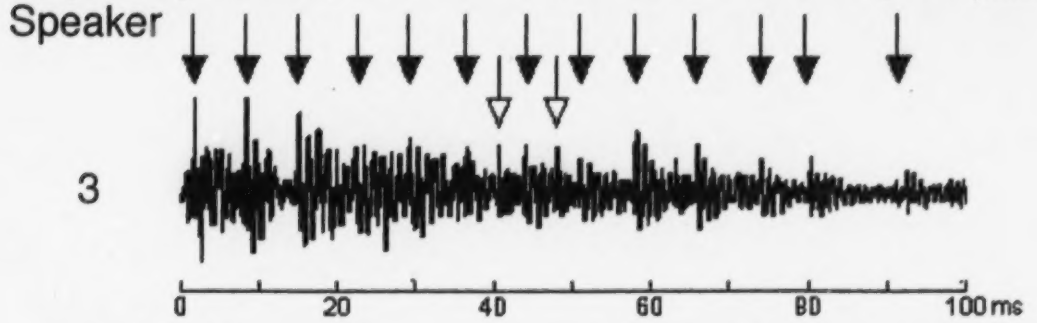



Figure 3. Solid arrows: the estimated onsets of the vocal pulses for speaker 3. Open arrows: other possible pulse onsets.

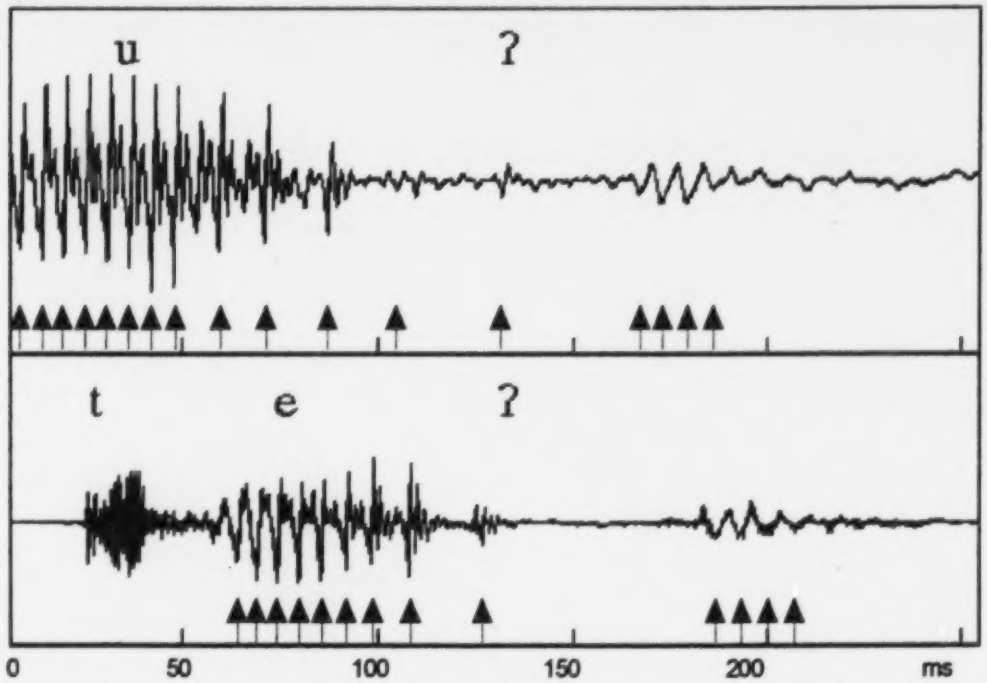
It is, of course, possible to calculate the degree of jitter (the standard deviation from the mean glottal pulse interval) for vowels with modal voicing. Doing this for the five Jalapa Mazatec speakers allows us to compare their modal and creaky vowels. For the speakers of Mazatec represented by the data shown in figure 2, creaky voice always had more jitter than modal voice. As you can see in Table 1, the modal voices had very regular pulses. The standard deviation from the mean interval between pulses was only .28 ms for the mean speaker. Putting this more informally, in modal voice, the intervals between 99% of all glottal pulses vary by less than a millisecond. (This is a rough interpretation of the statistics; many qualifications apply.) In creaky voiced vowels the standard deviation was always greater. It had a range from .58 to 5.07 ms, with a mean of 3.02 ms.

Table 1. The amount of jitter, measured as the standard deviation from the mean in ms, in the modal and creaky vowels of 5 speakers of Jalapa Mazatec.

| | Speaker 1 | Speaker 2 | Speaker 3 | Speaker 4 | Speaker 5 | Mean speaker |
|---|---|---|---|---|---|---|
| Creaky | .62 | .58 | 2.17 | 5.07 | 3.81 | 3.02 |
| Modal | .20 | .28 | .140 | .390 | .33 | .28 |

Creaky voice often occurs as the phonetic realization of a phonological glottal stop. Figure 4 shows the waveforms of parts of two Montana Salish words, [q'áq'łu?] "vein" and [kʷáte?] "quarter" (25c, a loan word). This is another case in which it is hard to say exactly when each pulse occurs. It is often unclear what should be counted as a separate glottal pulse. Nevertheless, what is apparent is that the glottal pulses occur at very irregular intervals.

Figure 4. The last parts of the Montana Salish words, [q'áq'ɬu?] "vein" and [kʷáte?]




"quarter (25c)". The arrows below the waveforms are attempts to mark each glottal pulse.

What we really want to know in order to describe different types of phonation is the shape of the glottal pulses produced by the vibrating vocal folds. Usually the vocal folds produce pulses of air as a result of being blown apart and then springing back together again. In modal voice their outward movement is a little slower than the rate at which they come together. Consequently the pulses of air that are produced are asymmetrical, with a slower rise and a sharper descent. Creaky voice and breathy voice produce different pulse shapes. Figure 5 shows a schematized version of the three different phonation types. When producing the creaky voice pulse at the top, the vocal folds come together very rapidly, making the closing phase much shorter than the opening phase. They then remain tightly shut, with a comparatively long closure. In a modal voice the closing phase is shorter than the opening phase, but the two are more nearly the same. In a really breathy voice, as schematized here at the bottom of the figure, the opening and closing phases are much the same length, and there is virtually no closure.

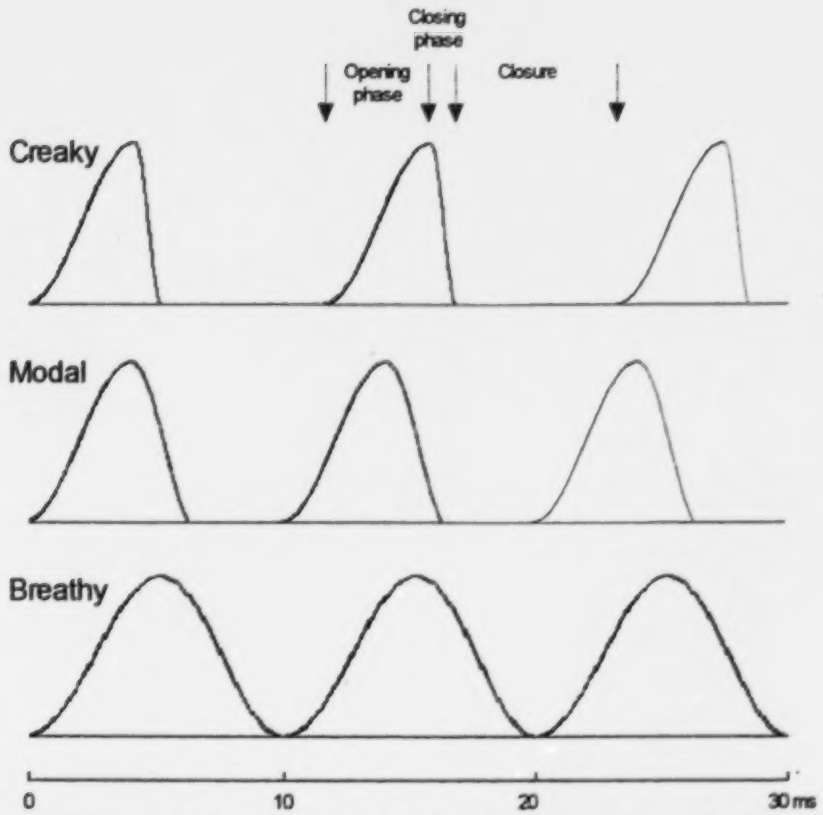



Figure 5. Schematized versions of the glottal airflow in three types of phonation.

The vocal fold pulses set the air in the vocal tract vibrating in ways that depend on the shapes of the cavities involved. The different vocal tract shapes produce the formant frequencies that characterize vowels and other sounds. The sound we hear — the waveform — is a combination of the vocal fold pulse and the resonances associated with the formants. If we want to see the vocal fold pulses, we must find some technique that can disentangle pulses from the rest of the waveform. This can be done by a technique known as inverse filtering, which removes the formant resonances, leaving the vocal fold pulse. It's a complicated technique that works best with comparatively steady state vowels and high quality recordings that include even the very low frequency variations in air pressure, all of which make it difficult to use in most linguistic investigations.

We can, however, infer something about the vocal pulse shape from the spectrum of a sound. The spectrum of a sound shows the relative energy in the harmonics. A creaky voice pulse, with a sharp closure, will contribute more to the high frequencies of a sound, and breathy voice vibrations will have much of the energy in the fundamental frequency (the first harmonic). Figure 6 shows narrow band spectra of three vowels in San Lucas Quiaviní, one of the Zapotec languages, spoken in Mexico. This language, like Jalapa Mazatec, contrasts three phonation types, creaky, modal

and breathy. As we have noted, in creaky voice the vocal folds snap shut quickly, producing a glottal pulse that has a great deal of energy in the higher frequencies; modal voice produces slightly less energy in these regions; and breathy voice has most energy in the fundamental frequency (the first harmonic). It follows that we can assess the extent to which a sound has breathy, modal, or creaky voice by measuring how much energy there is in the first harmonic in comparison with the energy in the higher frequencies

Figure 6 shows two ways of measuring the effects of the different vocal fold pulses, each involving a comparison with the amplitude of H1, the energy in the fundamental frequency. A dashed line has been drawn at the level of H1 in each of the spectra. The first method is to compare H1 and H2 (the second harmonic). In creaky voice H2 has almost the same amplitude as H1. In modal voice H2 is about 2.5 dB lower, and in breathy voice it is a little lower still, about 4 dB below H1. The second method is to compare the strongest harmonic in the first formant, here marked F1, with H1. In creaky voice H1 – F1 = –7.5 dB (the first formant is 7.5 dB above H1). In modal voice the first formant is about 4 dB above H1, and in breathy voice it is nearly 8 dB below H1.

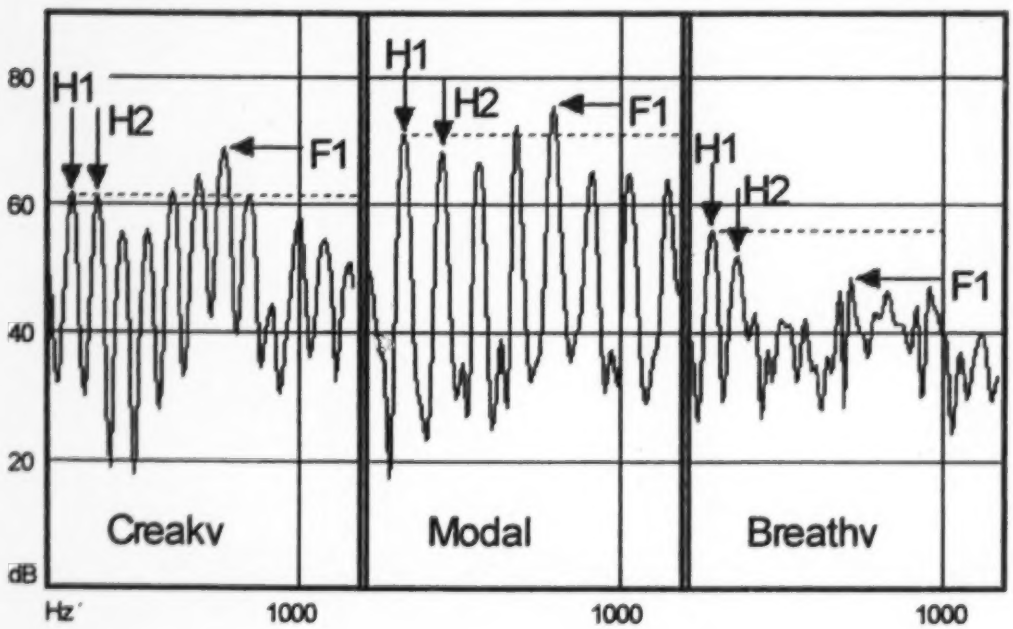



Figure 6. FFT spectra of creaky, modal and breathy **a** in San Lucas Quiaviní Zapotec. The dashed line marks the intensity of H1, the first harmonic, in each spectrum.

When these relationships are expressed graphically, as in figure 7, they can both be seen to reflect a tendency for the downward slope of the spectrum to increase as one

moves from creaky through modal to breathy voice. The breathier the voice, the greater difference between the first harmonic and higher frequencies.

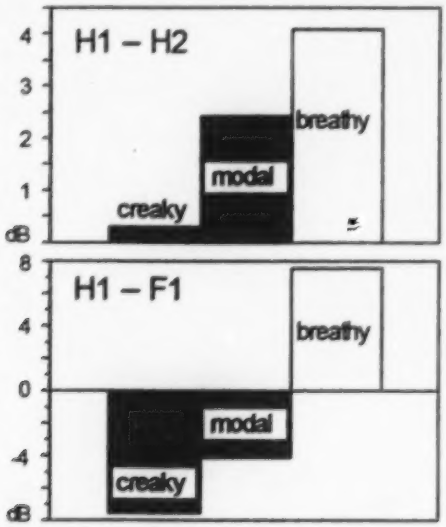



Figure 7. A comparison of two methods of measuring spectra of creaky, modal and breathy a in San Lucas Quiaviní Zapotec.

Finally, a note of warning must be sounded over the use of these spectral measures. Differences in phonation are not the only things that can affect the relative intensities of the harmonics. They are also affected by differences in pitch and in vowel quality. Consequently it is important to compare vowels that do not differ in quality (i.e. in the formant frequencies) or in pitch to any great extent. Subject to this proviso, measures of spectral slope are useful for specifying different voice qualities.

The perception of voice quality is a complex psycho-acoustic phenomenon. The relationship between the vibration of the vocal folds, the acoustic signal and listeners' perceptions is intricate and not well-understood, even by specialists. We recommend that you not rely upon any single acoustic measurement technique, and it is always useful to keep a record of what you hear and what you see in the wavefom or spectrogram. For a more in-depth discussion of the measurement techniques reviewed here, we suggest the following readings:

Ladefoged, P. (1996) *Elements of Acoustic Phonetics.* 2nd. Ed. Chicago: University of Chicago Press.
A thorough introduction to acoustic analysis techniques.

Laver, J. (1980). *The Phonetic Description of Voice Quality.* Cambridge: Cambridge University Press.
A description and analysis of different linguistic and pathological voice qualities.

Stevens, K.N. and H.M. Hanson. (1995) Classification of glottal vibration from acoustic measurements. In O. Fujimura and M. Hirano, eds., *Voice Quality Control.* Singular Publishing Group, Inc.: San Diego, CA. 147-170.
An explanation of spectral measurement techniques for voice quality analysis.